DOCUMENT RESUME

ED 225 347 FL 013 362

AUTHOR Lutjeharms, Madeline
TITLE Testing Reading Comprehension: An Example from German for Academic Purposes.
PUB DATE Aug 82
NOTE 13p.; Paper presented at the International Symposium on Language for Special Purposes (Eindhoven, The Netherlands, August 2-4, 1982).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Dutch; *German; Higher Education; Languages for Special Purposes; *Reading Comprehension; *Reading Tests; Second Language Instruction; Teacher Made Tests; Test Construction; Test Reliability

ABSTRACT

An approach to testing reading comprehension among Dutch speaking university students of German for academic purposes is described. The test items were designed to reflect the most typical comprehension errors made by speakers of Dutch when reading German academic discourse. The most frequent errors were due to vocabulary and declension difficulties. After reading a German language text, students are given open-ended questions about excerpts involving these problem areas. The questions require the students to provide loose translations of phrases from the text. Use of the first language for response eliminates the effect of productive skills. The open-ended items discriminate better than multiple choice items because the context is necessary for the right solution. The approach and variations have been tried several times with a group of 23 students, and the results compared and analyzed for reliability and validity. Reliability was found to be high, and correlations between detailed and global comprehension were low. Choice of test and item content is discussed. (RW)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

TESTING READING COMPREHENSION : AN EXAMPLE FROM GERMAN FOR ACADEMIC PURPOSES

ED225347

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Madeline Lutjeharms

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

Madeline LUTJEHARMS
Instituut voor Taalonderwijs
Vrije Universiteit Brussel

1. Introductory remarks about teaching methods[1]

1.1 Use of contrastive analysis

Source language Dutch and target language German being highly related, we make use of positive transfer in our teaching of LSP for students of sociology, political sciences, production management and communication sciences at the Free University of Brussels. We work with quite difficult texts from the start, "employing", as Fisiak (1981 : 8) expresses it, "the previous knowledge of the learner, informing him about similarities and differences...,also warning him about making false analogies and about the potential areas of interference". Syntax hardly contains any differences that present problems to the reader, and a large part of the vocabulary is recognizable. On the other hand, although we can presuppose a certain degree of specialized knowledge and a mastering of international terminology, in more abstract academic German texts there are some typical vocabulary problems. This in combination with the well-known features of academic discourse such as complex noun phrases and - typical only for academic texts from the humanities - a more complex syntax and long sentences, makes specific training necessary.

1.2. Exercises preparing and accompanying the reading

We make use of vocabulary exercises such as filling in blanks with or without guidance, aimed at problem words or word families. These may be either false friends or words where what Juhász (1970 : 92ff.) calls "lack of contrast" seems to come into play, words like "Ereignis, Ergebnis, Erfolg" etc., where Hungarian and Flemish students have the same problems. I would like to include in this group words like "Gewerkschaft, Grundsatz, Vorgehen, Verhalten, Verfassung, Verfahren" etc., where the individual morphemes are taken for granted, the words look familiar, but the meaning is not

FL 013 362

comprehensible without a dictionary or previous knowledge. Decoding errors may be due both to interlingual and intralingual inference. This type of words is very frequent in the kind of texts our students have to read.

I have set up exercises to give insight into morphological variants and lexical compounding, particularly with respect to regularities of word formation. Mastering of the process of nominalization, affixation and compound word formation helps to develop reading skills (Hoffmann 1976 : 258)

The modal verbs, which are partly false friends, are often encountered in their special, tense-linked uses (e.g. "Es dürfte/mag sein"). So exercises to these problems have been made. There are also transformation and perception exercises for the following grammatical phenomena : inflectional indicators, prepositions and certain structures like genitive attributes and nominalizations. For the inflectional indicators for example, students have to look for certain forms in a text or to answer questions on a text asked in such a way that the answers have to involve a change in declension from that given in the text.

The grammatical exercises are nearly always context-bound and we use authentic materials. We look for texts in which the relevant problem appears frequently. The vocabulary exercises are rarely context-bound; only so in specially written texts. In both kinds of exercises productive skills may be involved - in a very guided way -, for it is our experience that this supports reading, e.g. a quicker recognition of inflectional indicators, which are not always redundant and may easily be overlooked.

1.3. Exercises in global comprehension

The reading exercises themselves consist, for global text comprehension :

- of scanning or search reading with a specific information task
- of very general questions on the content - these may be in L1
- and of distinguishing the structure of the text and its information, which is very important for comprehension and especially for recall and retention of information (Van Dijk 1977 : 28).

Other exercises consist of allocating judgments on the text to the relevant parts of it and of underlining the main points of new information or opinion.

1.4. Exercises on detailed text comprehension

The students are taught detailed text comprehension via the analysis of complex sentence structures, the search for certain words or forms in the text, e.g. such as those expressing uncertainty, by questions on the contents, by "true or false" questions on the text and by transformation of nominalizations such as used for grammar.

2. Testing of reading comprehension

2.1. L2-difficulties in reading comprehension

I started testing reading comprehension by asking the students to make a summary of a German text in L1. This entailed evaluation problems. Did a subject leave a point out, because he or she considered it unimportant or because he or she couldn't decode it? Finally I only substracted points for evident comprehension errors. Most students grasped the outline of the text anyhow and for those who didn't it clearly was not a L2-problem.
Anyhow, this test procedure was not very satisfactory and quite time consuming as well. So I made an error analysis on summaries from several years and now a large part of the reading comprehension tests consists of items on frequent errors made by speakers of Dutch, when reading German academic discourse. These most frequent errors are mainly due to vocabulary and declension difficulties.

2.1.1. Vocabulary

- *false friends, among which conjunctions, modal verbs and interrogatives*
- *confounding words with or without "Umlaut" or overlooking it (e.g. fordern/ fördern, subjunctive 2)*
- *compounds where division into morphemes presents problems (e.g. "Amterhäufung", "Druckerzeugnisse")*
- *overlooking of separable prefixes of verbs*

2.2.2. Declension

- *overlooked inflectional indicator (e.g. "aller Interesse" read as "alle Interessen"; dative object read as accusative object, if in Dutch a preposition is required; subject after verb read as object[2])*
- *prepositioned feminine dative object misunderstood as a masculine subject[2]*

- confusion between plural and singular forms of nouns (plural without ending, singular of masculine "weak" nouns)
- confusion between articles and relative pronouns (e.g. "auf dem Programme auszuwiegen sind" read as "auf dem Programm" instead of "worauf")

Remarks from several colleagues indicate that these errors are not only typical for L1 speakers of Dutch.

2.2. Why use L1 ?

After the second world war most test development took place in the USA and L2-tests were constructed mainly for heterogenous L1 groups of testees. Besides, it was the period in which use of L1 was taboo in L2-teaching. So the source language certainly was not used in testing. In what Spolsky (1978/1980) calls the post-modern period in testing, which is our time, there is a recourse to procedures of the traditional period, however with many modifications and corrections. Among other procedures translation has become acceptable again and this not only for receptive skills (L2 into L1), but even for testing L2 oral or general language proficiency (Clark, 1980 : 18; Marton 1981 : 154,181)

As I work with homogenous L1 groups, utilize L1 in instruction both for transfer and as a teaching strategy to provide insight, use of the source language appears to me to be the most precise and efficient way to evaluate comprehension. Use of L1 enables the tester to eliminate productive skills or reading of sentences without context in the target language. I ask for very short translations. This might be problematic of course, as e.g. a translation factor might come into play. Translation has not been trained, as it is not a teaching aim. I have not yet done any research on this point. On the other hand there is the strong contrastive stamp of the items through which the reference to L1 is present anyhow.

2.3. The text for a reading comprehension test

2.3.1. Choice of text

The examination texts belong in general to the informing text type, are descriptive or argumentative and preferably written in a neutral, impersonal style. The texts could be classified into what Hoffmann (1976 : 185 ff) calls group C. Texts in a highly theoretical research style are too difficult for most of the testees because of their contents and so are not used for examination purposes. The choice of text is not so easy, since the text is

supposed to contain a more or less representative sample of comprehension difficulties. Quite often only a few difficulties appear with frequency, depending on the style of the author. Text lenghth amounts to about a thousand or slighltly more words.

2.3.2. An example

The test from which I give some examples was made for second year students of political sciences. The text is slightly more difficult than average. For this group of testees the result of the reading comprehension test counts for a quarter of the exam marks. They also have to prove some oral proficiency and to take a grammar test. About half of the students learned some German at school, the others had not studied German before coming to the university. They could attend a course of - in theory - 120 teaching hours. About half of the course time was devoted to reading exercises.

The text of this particular example consists of extracts from a whole page essay in the weekly "Die Zeit" by H. Jäckel, professor of political sciences in West -Berlin, entitled "Der "Radikalenerlass" (Die Zeit, 29.1.1982).

Some extracts from the test text :

... Dass dieses unter der grundgesetzlichen Ordnung durch einen ministeriellen Runderlass, wie er drei Wochen später auf den Behördenschreibtischen lag, überhaupt möglich sei, ist von der Mehrzahl der Kritiker offenbar ohne sonderliche Bedenken unterstellt und hingenommen worden. Widerspruch und Protest richteten sich nahezu ausschliesslich dagegen, dass die gemeinsam verabschiedeten "Grundsätze" angeblich gravierend in die Rechte und Rechtschancen der vielen einzelnen eingriffen, die als Beamte und Angestellteim öffentlichen Dienst von Bund, Ländern und Gemeinden tätig sind oder tätig sein möchten....

Dass öffentliche Meinung und meinungsbildendes Ereignis sich nicht immer decken, dass selbst ein zeitgenössisches Echo sich von den überprüfbaren historischen Fakten zuweilen ablöst und uneinholbar emanzipiert, ist freilich eine alte Erfahrung....

Sie [die Regierungschefs des Bundes und der Länder] haben zunächst
15 einmal leicht verkürzt zwei beamtengesetzliche Bestimmungen zu der das öffentliche Dienstrecht kennzeichnenden besonderen Treuepficht zitiert, und sie haben in einem Nachsatz zutreffend festgestellt, was von keiner Seite in Zweifel gezogen war : "Es handelt sich hierbei um zwingende Vorschriften"...

The typical characteristics of academic discourse[3] in this text are especially the multiplication of pre-headword modifying strings (line 15-16), complex prepositional phrases and prepositioned subject clauses (line 1-**1**, 10-13, cf. Birkenfeld, 1981 : 79). There are rather few semantically weak verbs (if we don't include "sein", which Hoffmann, 1976 : 359 nevertheless does) and consequently hardly any nouns ending in -ung, -heit, -keit etc. the continuous passive (line 3-4) and its paraphrases are represented. The terminology is very appropriate for the target group, although there are no complex compounds and derivations. Words like "beamtengesetzlich" (line 15) or "freiheitsbeschränkend" are no problem.

The students read the text at first for about 10 to 15 minutes, then they are allowed to use a dictionary for 20 minutes. Only after this are they given the test items. To answer these, with the text, but without dictionary, they have about 35 minutes, which puts them under pressure. The whole procedure has been tried out several times during teaching time and at home with tests from previous years.

2.4. Questions on detailed text comprehension

2.4.1. Examples of test items

There are 20 items altogether.

item number		difficulty (p)	discrimination index
3.	line 3: translate "ist von der Mehrheit der Kritiker offenbar".	.63	.55
	problems : von (false friend in passive voice), Kritiker (plural), offenbar (false friend, following context helps)		
4.	line 5 - 6 , translate "dass die gemeinsam verabschiedeten Grundsätze".	.39	.49
	problems : "Grundsatz" often confounded with "Grundgesetz", "verabschiedet" has to be guessed from context or looked up.		
5.	line 9 : translate "tätig sein möchten".	.52	.65
	problem : modal verb (false friend)		

8. line 10 : translate "meinungsbildendes Ereignis"
problem : "Ereignis" difficult but often repeated word
.76 .37

10. line 15 : to what word does "der" refer to or belong to ?
.43 .55
problem : multiplication of pre-head word modifying strings.

The testees know that they don't have to give absolutely literal translations, approximations are sufficient, as long as content is reproduced correctly. Lubricators, for example, since they have relatively little significance, may be freely translated or omitted. Omission may even be the best solution.

2.4.2. Why mainly open-ended items?

When carrying out an experiment with comparable items I found that multiple choice items in such a test are somewhat less discriminating than open-ended items and furthermore considerably more difficult (for more details : Lutjeharms, forthcoming; I also considered a non-random cloze). The construction of MC-items is difficult and time-consuming. Open-ended items demand more correction time, but with relatively small groups of subjects the ease of correction of MC-items does not outweigh their disadvantages. MC-items are useful in that they can exclude solutions that are perhaps plausible in context, but less precise or good. This is the only case in which I still use them. The experiment, which has been repeated, also showed that the correlations between test score and attendance in class are higher for open-ended questions than for MC-items. This probably shows that a MC-item allows more intelligent guessing than we like to admit. Now I have no objection to intelligent guessing - on the contrary I stimulate it - but as this test is meant to be a criterion-referenced achievement test, that is not what I am interested in in the first place.

2.4.3. Item content

What skills and knowledge are demanded from the testees with such items? On the one hand lexical and grammatical knowledge they have learned during the course, but on the other hand the application of this knowledge in the right way. Gaps may quite often be compensated by help from the context.

Test items seldom exceed the level of sentence comprehension, even though the context may be most efficiently exploited as an aid to comprehension and to exclude errors, if one pays attention to internal logic of the paragraph or the text as a whole. In the test presented here there is one bad item with a negative discrimination index (-.06), an item on a false friend, where context wasn't really helpful and its content was not relevant for the text either. It appeared to be one of the most difficult items as well ($p = .33$)[4]. Similarly for the only other really bad item in the 1982 reading comprehension tests, an easy item on part of an enumeration, context was no help. This shows how important context is, when writing such items.

2.4.4. Reliability and discrimination indices.

The reliability coefficient of this test, .87 (Cronbach's alpha), is rather high, considering the small number of testees ($N = 23$) and the fact, that the test had not been tried out beforehand, to throw out bad items. For a try-out I lack a comparable group of testees. However test analysis makes it possible to establish what item types or forms are not very discriminative or even have negative correlations (cf. Lutjeharms, 1981 : 92 f.)

The reason for a high reliability might be that the items test more or less the same thing, which would be at the cost of validity, as reading comprehension is a very complex process. These items contain different types of vocabulary and grammatical problems, but they emphasize especially the difficulties that might interfere with correct decoding of the text. Reading comprehension certainly involves much more, for instance finding main ideas or essential information, or seeing relations between these etc. In fact with the items on detailed text comprehension I only test the preconditions for the global text comprehension.

2.5. Questions on global text comprehension.

That is why there have to be some questions on global text comprehension. As there are only 3 to 5 of these questions - it is not possible to make more questions, for they would not be global anymore -, calculations on reliability are pointless. This means correlations with the global text-comprehension results have to be used with some caution.

I try to find very global questions about the general attitude or opinion of the author, about facts that are dealt with repeatedly or about the structure of the text. Answers should be very short. All is in L1. I don't go beyond text comprehension. There are no questions about recognizing a writer's purpose, tone and mood, or identifying a writer's techniques. Here I am quoting those points of the Davis'scheme for reading comprehension tests, that I don't use in my test (Maas-De Brouwer, undated : 8). Subjects know this kind of style and text type anyhow. It seems to me that these points of the Davis'scheme apply less to language for academic purposes.

2.6. Relationship between global anddetailed text comprehension.

Correlations between detailed and global comprehension are rather low. I had expected this, as attending class is important for the detailed comprehension items, whereas in the global comprehension questions intelligence and knowledge of the world play a very important (the most important?) part, especially as a dictionary can be used and positive transfer from L1 is important.

Correlations (Pearson)

Group	Production management			Communication sciences			Political sciences		
Year	'80	'81	'82	'80	'81	'82	'80	'81	'82
N	23	32	23	49	80	69	33	-	23
Reliability (Cronbach's alpha)	.80	.86	.90	.81	.90	.85	.81	-	.87
Detailed comprehension/ grammar test	.62	.58	.60	.44	.72	.48	.54	-	.73
D.C./oral proficiency	.74	.69	.68	-	-	-	.54	-	.61
D.C./global comprehension	.52	.59	.46	.39	.50	.44	.46	-	.52

When I compare the detailed comprehension scores with the score on the grammar test (20 items, questions on the reading comprehension test, whereby change in declension is necessary for the answer), with the marks for oral proficiency (classic interview, global judgment by 2 raters) and with the results of the global comprehension questions, the correlation between the two parts of the reading comprehension test has always been the lowest one.

This proves I test here two quite different aspects of reading comprehension. We spend a lot of time exercising and trying to improve global text comprehension during the course. However if I compare a student's result in the global reading comprehension test part with his or her attendance in class, I can only conclude that the last factor has little influence on results for this part of the test.

Notes

1. *I should like to thank my colleague Frank Winter for reading through the English text.*

2. *The tendency to inversion is typical for language for academic purposes; cf. Hoffmann (1976 : 366) : "Auffällig ist im Russischen und Deutschen auch die Tendenz zur Inversion, die bis zum Vordringen des Objekts aus seiner überwiegenden Schlussstellung an den Satzanfang führen kann". In such cases the inflectional indicators can be very important, for instance the adjective ending in a feminine dative object.*

3. *cf. Wieser, 1978 for German; Hoffmann, 1976 : 240, 276, 358, 239 etc. gives most of these characteristics as typical for language for academic purposes in general - Russian, French, English, German - and especially for German adds compound nouns and a tendency to inversion of objects (257, 366).*

4. *The item is : translate "mit dieser Auskunft", from the sentence "In der Tat machen die Lexikon-Autoren es sich mit dieser Auskunft einfach : Sie nehmen auf und teilen mit, was.... ", often translated as "result" ("uitkomst") instead of "information".*

References

BIRKENFELD, H. (1981), G. Gutterer, B. Latour : Grammatik in wissenschaftlichen Texten (book review). Fachsprache, 3 (2), 79-83

CLARK, J.L.D. (1980) Toward a Common Measure of Speaking Proficiency. In J.R. Frith (ed.) Measuring Spoken Language Proficiency. Washington : Georgetown U.P.

FISIAK, J. (1981) Some introductory notes concerning contrastive linguistics. In J. Fisiak (ed.) Contrastive Linguistics and the Language Teacher. Oxford etc. : Pergamon Press, 1-11

HOFFMANN, L. (1976) Kommunikationsmittel Fachsprache. Berlin : Akademie-Verlag.

JUHÁSZ, J. (1970) Probleme der Interferenz. Budapest, Akadémiai Kiadó.

LUTJEHARMS, M. (1981) Zum Testen des fachorientierten Leseverständnisses : ein Beispiel. In T. Culhane, C. Klein-Braley, D.K. Stevenson (eds.) Practice and problems in language testing 4. University of Essex, Occasional Papers no. 26, 88-94.

LUTJEHARMS, M. (forthcoming) Leseverständnis testen im Bereich DaF : einige Überlegungen zu Testinhalt und Progeverfahren. In C. Klein-Braley, D.K. Stevenson (eds). Practice and problems in language testing 2. Frankfurt etc. : Verlag Peter D. Lang. Orbis Linguisticus 3.

MAAS-DE BROUWER, T.A., SAMSON-SLUITER, D.M.M. et al. (undated). Verslag project 715 : Validatieonderzoek tekstbegriptoetsen moderne vreemde talen, fase 1. Arnhem : CITO. CITO-MEMO 332.

MARTON, W. (1981) Contrastive analysis in the classroom/Pedagogical implications of contrastive studies / Some more remarks on the pedagogical use of contrastive studies. In J. Fisiak (ed.) Contrastive Linguistics and the Language Teacher. Oxford etc. : Pergamon Press, 147-183.

SPOLSKY, B. (1978) Language testing : art or science? Main lecture presented at the AILA World Congress, Stuttgart 1975. In G. Nickel (ed.) Language Testing. Stuttgart : Hochschulverlag, 215-234.

SPOLSKY, B. (1980) Some ethical question about language testing. In C. Klein-Braley, D.K. Stevenson (eds.) Practice and problems in language testing 1. Frankfurt etc. : Verlag Peter D. Lang, 1-17.

VAN DIJK, T.A. (1977) Semantic Macro-Structures and Knowledge Frames in Discourse Comprehension. In M.A. Just, P.A. Carpenter (eds.) Cognitive processes in comprehension. Hillsdale, New Jersey : L. Erlbaum, 3-32.

WIESER, J. (1978) Allgemeines Deutsch der Fachsprachen. Eine kurze Charakteristik. In H. von Faber, K. Echtermeyer (eds.) Werkstattgespräch Kafro : Fachsprachen. München : Goethe Institut, 12-42.